DOCUMENT RESUME

ED 061 915 24 HE 003 040

AUTHOR Fernald, Peter S.; DuNann, Deborah H.
TITLE Investigation of Individualized Instruction for Large College Classes. Final Report.
INSTITUTION New Hampshire Univ., Durham. Dept. of Psychology.
SPONS AGENCY Office of Education (DHEW), Washington, D.C. Bureau of Research.
BUREAU NO BR-0-A-068
PUB DATE Mar 72
GRANT OEG-1-71-0004
NOTE 45p.

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Higher Education; *Individual Instruction; *Individualized Instruction; *Student Centered Curriculum; Student Development; *Teaching Methods

ABSTRACT

In the present study, a form of individualized instruction is compared with the conventional lecture approach. Hourly and final exams given to both groups indicate a superiority for students under individualized instruction, which, while highly significant statistically, represents only moderate pragmatic gains. Attitude questionnaires indicate more positive attitudes for students receiving individualized instruction; however, it appears that individualized instruction produces no study skills that the students later employ under the conventional system, and it is no more beneficial for low performing students than for high performing students. It was found, in addition, that those receiving individualized instruction are able to evaluate their mastery of the assigned material more accurately than students taught by the lecture method. It is recommended that future research be directed at separating some of the many factors that make up individualized instruction so as to identify their relative contributions to this method of instruction. (Author/HS)

FINAL REPORT
Project No. OA068
Grant No. OEG-1-71-0004

INVESTIGATION OF INDIVIDUALIZED INSTRUCTION FOR LARGE COLLEGE CLASSES

Peter S. Fernald and Deborah H. DuNann

Department of Psychology

University of New Hampshire

Durham, N. H. 03824

March 1972

U. S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

Final Report

Project No. 0A068
Grant No. OEG-1-71-0004

Investigation of Individualized Instruction for Large College Classes

Peter S. Fernald and Deborah H. Du Nann

University of New Hampshire

Durham, N.H.

October 1971

The research reported herein was performed pursuant to a grant with the Office of Education, U. S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgement in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

Contents

ABSTRACT

A form of individualized instruction was compared with the conventional lecture approach. Hourly and final exams given to both groups indicated a superiority for students under individualized instruction, which, while highly significant statistically, represented only moderate pragmatic gains. Attitude questionnaires indicated more positive attitudes for students receiving individualized instruction. Individualized instruction produced no study skills which the students later employed under the conventional instructional system, and it was no more beneficial for low - than for high - performing students. Those receiving individualized instruction were able to evaluate their mastery of the assigned material more accurately than students taught by the lecture method.

Conclusions from this study are that individualized instruction is of immediate and moderate value in helping students in large classes learn subject matter, but implementation of the method should follow only after careful analysis of overall costs and gains. It is recommended that future research be directed at separating some of the many factors which make up individualized instruction so as to identify their relative contributions to this method of instruction.

INTRODUCTION

The use of large lecture sections at the college level has been rapidly growing as a result of increasing societal demands for a college education, "open admissions" policies of many state and junior colleges, and the worsening financial situations in most institutions of higher education. The many problems accruing with large class instruction have been clearly presented (Jensen, 1969; Sanford, 1909, cited by McKeachie, 1968) and are obvious to most teachers as a result of experience. The fact that the use of large lecture sections is an unattractive solution to increasing student-faculty ratios has been recognized for years among professional educators (Day, 1966, cited by Jensen, 1969) and more recently in lay literature (Time, 1971).

New technological advances have led some teachers to implement changes involving automation, for example, closed circuit television, programmed instruction, the use of computers in teaching and grading, and other similar approaches. Aside from the fact that research has failed to support the superiority of such aids over more traditional methods for teaching purposes at the college level (Roderick & Anderson, 1968; McKeachie, 1968), it is clear that automated instruction can increase the problems associated with a lack of personal interaction with the instructor. Not only are most of these materials expensive, but students often indicate feelings of being treated more like numbers than people (Volin, 1967). One educator specifically notes that the growth of already present "student-machine relation[s]...would be a poor substitute for direct social interaction" (Keller, 1968, p. 87).

Individualized Instruction

In an effort to circumvent the problems of depersonalization inherent in large class instruction and technological aids, several instructors have recently described a new approach which utilizes the enthusiasm and talents of undergraduate students for teaching purposes (Keller, 1968; Ferster, 1968; Malott & Svinicki, 1968; Domjam & Du Nann, 1969). While the methods and objectives of their courses differed in some ways, the courses were similar in 1) the division of course material into small units (typically chapters) which were followed by frequent evaluations and immediate feedback; 2) the structuring of the course such that one unit could not be attempted before the previous unit was successfully mastered; 3) the opportunity for greater self-scheduling on the part of the student by providing him many chances to satisfy criterion performance (in some courses he was even allowed to progress through the course at a completely self-chosen rate) and 4) the greatly increased opportunity for individual contact with an instructor so that the student could ask questions, express ideas, engage in small group discussion, and receive help with special problems.

Certain differences between the courses are notable, especially the frequency with which evaluations were instituted, and the form in which evaluations were made. Ferster (1968), for example, used self-scheduled interviews which were followed by less frequent essay quizzes; Keller (1968) used self-scheduled fill-in and short answer quizzes; and

Malott and Svinicki used daily quizzes composed of "construction type items* along with student-led four-man discussion groups which were peer graded. However, the procedures are sufficiently similar to one another along the four dimensions listed above, and different enough from the conventional large lecture procedure to have received a generic term "individualized instruction."

This kind of course is appropriately reminiscent of programmed learning--both have emanated from the operant conditioning model--but the provision for social interaction in individualized instruction makes it somewhat different from programmed instruction. Keller has expressed the relationship between this method of teaching and programmed instruction:

> There is the same stress upon analysis of the task, the same concern with terminal performance, the same opportunity for individualized progression, and so on. But the sphere of action here is different... The response is not simply the completion of a prepared statement through the insertion of a word or phrase. Rather, it may be thought of as the resultant of many such responses, better described as the understanding of a principle, a formula, or a concept, or as the ability to use an experimental technique. Advance within the program depends on something more than the appearance of a confirming word or the presentation of a new frame; it involves a personal interaction between a student and his peer, or his better, in what may be a lively verbal interchange, of interest and importance to each participant (1968, pp. 84-85).

This kind of course is seen as contrasting sharply with the typical large lecture section in which 1) course content is divided up into large units, typically five to 10 chapters, which are followed by hourly exams, typically two or three a semester with a final exam; feedback as to performance on these examinations is delayed by at least 24 hours and sometimes up to one or two weeks; 2) a student can perform poorly on all the course content, there being no regulation that he successfully master one unit before he attempts another; 3) there is no opportunity for student self-scheduling: all students are expected to perform on examinations at the same hour; and 4) there is little opportunity for individual interaction with an instructor: questions in class are handled as quickly as possible, ideas and feelings are hardly dealt with at all, and students receive special help only if they are persistent enough to arrange it with a graduate student, who typically has little contact with the lectures, and even less with the evaluations.

Individualized Instruction vs. Conventional Teaching Methods

Several studies have compared individualized instruction with conventional lecture methods (Keller, 1968; McMichael & Corey, 1969; Sheppard & MacDermot, 1970; Johnston & Pennypacker, 1971; Born, Gledhill, & Davis, 1971). Although these studies generally support the superiority of the individualized approach, there are notable variations in research designs and problems in interpretation.

*left undefined in their 1968 paper.

For example, one problem results from the confounding of task difficulty with instructional method (Johnston & Pennypacker, 1971). The fact that it is easier to remember material for frequent, self-scheduled quizzes than to retain larger chunks of material for longer but more infrequent examinations was recognized by Malott & Svinicki (1968) when they noted that review quizzes might be necessary to demonstrate superior performance on final examinations with individualized instruction. However Johnston & Pennypacker concluded that individualized instruction was superior because students in the course produced better response rates (more accurate and faster) on short, frequent quizzes, than did control groups responding to the same items on infrequent hourly examinations.*

A somewhat more convincing study (McMichael & Corey, 1970) compared end-of-the-semester performance of students enrolled in a Keller-type course with the performance of students in three control sections employing the conventional lecture format. Performance was measured with a 50 item multiple choice examination and revealed a clear superiority of the individualized section which had an average of six points higher than the control groups. Unfortunately, however, as the investigators point out, they did not control for teacher variables, having had different instructors teach the various groups (McMichael and Corey, 1970). Although indirect evidence was presented to suggest otherwise, it is possible that the instructor of the experimental group was a better teacher, or that he simply displayed more energy, enthusiasm, and invested more time.

In an effort to provide a more adequate control over teacher variables, another study (Sheppard & MacDermot, 1970) compared end-of-the-semester performance of students in an individualized section against that of students in a conventional section. All students heard the same lectures, were exposed to the same study questions, and spent the same amount of time in class. A bias in favor of the control group was introduced by having 50% of the final grades in this section determined by performance on the exam, while grades in the individualized section were not influenced by examination performance. Yet, test scores were superior in the individualized group. This is the only study to the author's knowledge which reports an attempt to control for teacher variables.

The above study is, however, comparable to several others in demonstrating that the effects of individualized instruction generalize across types of test items. In this case, superior performance of students in the individualized condition was shown on both essay and multiple choice items, even though the kinds of responses required in the interview situations were much more analogous to the essay type question.

*These authors did compare performance on end-of-the-semester essay examinations, and concluded that individualized instruction was successful because it did not hinder performance on this measure. However, this evidence is not at all convincing in light of the fact that several studies (to be discussed shortly) demonstrate that when superiority in performance does occur, it generalized across different kinds of test items.

Such generalization of the effect across types of items was also reported by others (Alba and Pennypacker, 1971) who administered pre- and post-tests consisting of fill-in and multiple choice items to students in a conventional group and students in an individualized course. The frequent evaluations were made with fill-in and short essay questions, and the superior performance of students in the individualized section generalized on end-of-the-semester multiple choice measures.

Similarly, another study conducted with individualized instruction (Born, Gledhill, and Davis, 1971) found superior performance on essay, fill-in, and multiple choice items given at the end of the semester. This study was particularly interesting because it was the first to have students exposed to more than one condition. The researchers employed a design which utilized a Keller-type section, a modified Keller-type section (in which the student determined the size of the unit he would be quizzed on), a lecture section, and a section which rotated through the three conditions. The authors found performance superior in both forms of the individualized course on essay, fill-in, and multiple choice items given at the end of the semester. Analysis of the performances of students in the rotating section indicated that individualized instruction had its major effect on students with average to poor academic records.

However, a shortcoming of this study was that course content was confounded in the rotating section. All students in this section started with the lecture condition and then went to the individualized instruction groups. Since different material was covered in these different conditions, it is possible that performance improved during the latter part of the semester because the material was easier, not because the individualized instruction method was better. While such an upward trend in performance was not observed in the upper half of the rotating section, or in the individualized groups, this may have been due to a ceiling effect for the former (i.e. the good students have less room to move up) and a confounding of content with method for the latter. Though these are unlikely possibilities, they could be eliminated with a design which manipulated method while keeping course content constant. Since most instructors prefer to follow a particular sequence such as presenting basic processes before more molar concerns, solutions to this problem could be achieved by having one group move from a lecture condition to an individualized condition, and another move from an individualized condition to a lecture condition. The present research employed such a manipulation. Teacher variables were controlled by having the same teacher lecture both individualized and conventional sections, and content variables were controlled by having the sections rotate in both directions. It was felt that this design would provide the most stringent empirical test of individualized instruction to date.

The Differential Effect of Individualized Instruction on Strong and Weak Students

The suggestion by Born et al. that individualized instruction helps the weaker student more than it does the stronger student deserves more

attention, as the practical implications of such a finding are important. While it would seem reasonable to expect that performance would be differentially effected by individualized instruction, the data cited by these authors do not necessarily support the notion. While it was shown that the five students in the lower half of the class (as determined by the first examination scores) improved a greater amount of the second examination (after individualized instruction) than did the five students in the upper half of the class, such a finding is logical in terms of 1) a ceiling effect for the upper half of the class; 2) differential motivation, such that lower students try harder to improve than do better students.

Thus, the only way to meaningfully evaluate the differential effects of individualized instruction is to compare its effects for good and poor students against a baseline provided by comparable students in conventional conditions. Use of the crossover design in this study permitted such an analysis by the comparison of students who changed from conventional to individualized conditions to students who remained in the conventional condition throughout the semester.

Individualized Instruction and the Acquisition of Study Skills

Use of a design in which students move from the individualized instruction condition to the lecture condition enables one to answer another important question. Because students in the individualized instruction condition are encouraged to regularly study the course material and become aware of what they know and what they don't know through frequent feedback before examinations, perhaps they are actually acquiring better study skills. Would such improved study habits transfer to a condition in which the student is again faced with the conventional lecture format?

A number of studies conducted over the last forty years support claims that academic performance is improved by courses at the college level specifically designed to improve study skills (cf. Entwistle, 1960; Di Lorenzo, 1964), although those variables which are most important in this kind of training have not been clearly outlined (Rentel, 1966). While there are no reports to the authors' knowledge of any studies concerned with content courses which simultaneously train study skills, a study on retention and transfer as a function of feedback bears mention (Sassaranth & Gaverick, 1965). Four levels of feedback were given to different groups of students after three examinations: a) scores were announced without information as correct and incorrect answers; b) correct answers were written on the borad; c) page numbers of items were written on the board so students could look up wrong answers, and d) items were discussed by the instructor. Retention was measured by scores on final examination items which were the same as the items on the first three examinations, and transfer was measured by scores on final exam items which were different than those on the first three exams, but which covered the same material. On both the retention and transfer tests, the group which had the questions discussed in class performed best, and the group which had only scores read performed worst, with the other two conditions in between. The authors note that the results are congruent with earlier research on types of feedback (Stone, 1955) and with the explanation that the value of feedback is directly

proportional to the amount of information contained in it.

In the case of the present research, a similar test for transfer was made. One group received extensive feedback with discussion, while another group received little. Thus information level was varied in that in the individualized group, both a greater amount of information was given (i.e., a greater number of test items were discussed) and a greater depth of information was given (i.e., more discussion was allowed regarding the correct and incorrect answers). The primary difference between the present study and that of Sassaranth and Carverick's was that the present one tested transfer on new material, while the previous study tested transfer on the same material, but with different test items.

Ability to Evaluate Mastery

Disappointed and frustrated, failing students offer a variety of rationalizations for their low performance: "I knew the material but the exam didn't cover the important points"; "I could do much better on an essay exam"; "I can't understand it, I tutored my friend and he got a higher grade than me." Other less defensive students say, "I thought I understood the material, but it seems I didn't".

Apparently some students accurately evaluate their mastery of subject matter while others do not. What is the difference between the two? One possibility is that the accurate evaluator, unlike the inaccurate one, asks himself questions and seeks answers as he studies. We might say he had learned how to test himself; he becomes his own feedback monitor.

It would appear that teaching the student to become his own feedback agent is a very important aspect of learning, yet education researchers have devoted little, if any, attention to this topic. Can the ability to evaluate one's own progress in learning be taught? This question relates to the ability to evaluate mastery hypothesis which states that the ability to assess mastery will improve through experience in the individualized instruction condition.

Four Hypotheses

The present study tested the following hypotheses:

H_1 (Method) Material covered under the individualized instruction method will be better mastered as measured by multiple choice examinations than material covered under a conventional large lecture section method.

H_2 (Differential Help) Individualized instruction will help the weaker student more than it does the stronger student, as measured by change in scores from the first examination to the second.

H_3 (Transfer) Study skills employed under the individualized instruction method will transfer to situations in which the student is again

faced with the conventional large lecture format, and performance will subsequently improve.

H_4 (Evaluation of Mastery) Students in individualized instruction will learn to evaluate their mastery of course material.

METHOD

Subjects

Subjects were University of New Hampshire undergraduates enrolled in two sections of Introductory Psychology, each section containing approximately 280 students. All subjects were assigned the same reading, and took the same examinations. Students in all conditions were told that they had enrolled in an "experimental" course where new teaching in the "experimental" condition."

Individualized and Conventional Conditions

In the conventional (C) condition, students attended three 50 minute classes each week. These classes followed the conventional procedure, in that each meeting was devoted primarily to lecture, with occasional demonstrations and movies, and the groups attending these classes were large (280 or 140 students). Lectures given to the conventional group were designed to follow the textbook material closely, an effort being made to equate the material to which both individualized and conventional groups were exposed.*

In the individualized (I) condition, students attended one class each week which was devoted entirely to lecture. This lecture was given by the same instructor who conducted the lectures for the conventional group, and much care was given to make this lecture as identical as possible as the first one given each week to the conventional group. In addition, students in the I condition were assigned in groups of approximately 15 to an undergraduate student assistant who administered a brief quiz of 10 items over one chapter each week, under a "Doomsday Contingency" (Malott & Svinicki, 1968). This contingency refers to the regulation that within four sessions, each student either must pass at least one quiz with 8 out 10 or drop the course.** The purpose of this regulation was

*Because the instructor was required to equate material covered in both conditions, his lack of freedom to discuss issues or bring in new ideas not covered in the text, can be considered a departure from the conventional classroom situation. On the other hand, because most lecturers feel compelled to get through a preplanned lecture, and because discussion in large groups is usually quite minimal, these constraints also exist to some extent in the conventional class.

**While this contingency was introduced to the students as part of the course procedure, no one was actually forced to leave the course. Requirements were eased to allow students who obtained 7 out of 10 on the fourth quiz and had taken 3 previous quizzes to remain in the course. Reasons for not actually implementing the Doomsday Contingency are presented in the discussion.

to enhance motivation on the part of the student to keep up with the course.

After each quiz, the student assistant conducted a discussion of correct and incorrect answers and of any other material the students wished to consider. Occasionally demonstrations were conducted which were highly similar to the ones performed for the conventional group. Student assistants provided tutoring and individual help to students who had special difficulties in meeting the weekly requirement.

Designs

Subjects were assigned to four groups of approximately 140 each:

1. II group started the semester under individualized instruction and remained under individualized instruction throughout the entire semester.

2. IC group started the semester under the individualized condition and changed to the conventional condition after the mid-semester examination.

3. CI group started the semester under the conventional condition and changed to the individualized condition after the mid-semester examination.

4. CC group started the semester under the conventional condition and remained under the conventional condition throughout the semester.

Three measures of performance were taken: the first and second hourly exams, of 45 items each, each covering one half of the course material, and the final exam, which was 90 items, and covered material on both hourly exams, plus an additional chapter which was assigned during the independent reading period conducted at the university. This period extended for two weeks between the end of the semester and the final exam period. Thus, three measures were available for each of the four groups; they are depicted in Figure 1.

Figure 1

Group	N	Hourly 1	Hourly 2	Final
II	140	A	E	I
IC	140	B	F	J
CI	140	C	G	K
CC	140	D	H	L

For example, scores in Cell B consist of first hourly examination scores for the group which started with individualized instruction, and went to the conventional condition after the first examination. Likewise, Cell A consists of scores on the first hourly obtained by the group which

started individualized instruction and remained there during the entire semester. Thus at the time of the first hourly, these groups were the same, and they were combined in certain analyses, while after the first hourly they were treated as different groups. The same applies to Cells C and D. Tests of the four hypotheses are discussed with reference to Figure 1. In addition, other designs were used to test the same hypotheses, and they will be introduced as the design for each hypothesis presented.

Hypothesis$_1$ (Method)

The hypothesis was that material covered under the individualized condition would be better mastered than the same material covered under the conventional condition. The main dependent variable used to test this hypothesis was final examination scores, since it is usually terminal behavior which is considered to be the most important by teachers. However, because final exam scores do not reflect which material was learned best, scores on the first and second hourly exams were analyzed as a check against findings derived from final exam scores.

The design used to analyze final exam scores is pictorially represented in Figure 2.

Figure 2

Factor B: condition during second half of semester	Factor A: condition during first half of semester	
	II	CI
	IC	CC

In this design, Factor A is what happened to the student in the first half of the semester (individualized or conventional condition) and Factor B is what happened to the student during the second half of the semester. If the method has an effect during both halves of the semester, main effects for each variable should be observed. The cells which are of most importance are II and CC which should be significantly different from each other, II being greater than CC. IC and CI should fall somewhere in between them, since these groups had only half a semester of individualized instruction. The prediction is that CC < CI or IC < II, if there is no positive transfer from individualized instruction. Any transfer effects would bring up the IC group so that it would approach II group. Because the position of the IC group depends on the effects of transfer, then, the prediction for the method hypothesis is that CC < CI < II, and IC group should not be less than CI or more than II.

The second test of the method hypothesis was achieved by comparing the effects of individualized instruction on the first examination. All students having had individualized instruction were compared with all students having had conventional instruction, and the prediction, in terms of Figure 1, was that Cells A + B > C + D.

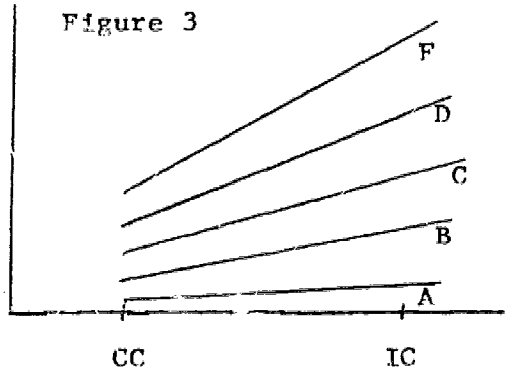
The third test of the hypothesis involved comparing scores on the second hourly exam with the prediction that CC < II. Because CI and

IC groups had by this time been involved in a change of conditions, they were dropped from the analysis because of possible confounding effects of change.

Hypothesis$_2$ (Differential Help)

The hypothesis was that individualized instruction helps poorer students more than it does better students. As noted before, in order to adequately test this hypothesis, it is necessary to compare improvement of the IC group against a baseline provided by the CC group. A 2 x 5 analysis of variance provided the design, where the two level factor refers to treatment conditions (CI and CC) and the five level factor refers to ability levels, arbitrarily defined in terms of grades received on the first examination. The dependent variable was difference scores from the first improvement (or deterioration, as the case may be). In order for the hypothesis of differential help to be confirmed, an interaction between these two factors must be found such that the difference between improvement scores for the CC vs. CI groups must be greater for poorer students than it is for stronger students. This interaction is pictorially illustrated in Figure 3. In order for the hypothesis to be confirmed, the slopes of the lines must uniformly increase as shown.

Figure 3



Because of differential motivation for students who received F's vs. those who received A's on the first exam, a main effect of the grade factor is predicted such that F students should change more than A students; hence the lines should stack up with F being higher than D, D being higher than C, and so on. Because of the effects of individualized instruction, it is expected that the lines would always be slanting in the same direction, such that CI points always fall above CC points for each line. However, if the method helps lower students moretthan it does better students, one would expect that the lines would slope differentially, as drawn. Thus two main effects and an interaction are predicted, the interaction serving as a test of the hypothesis of differential help.

Hypothesis$_3$ (Transfer)

The hypothesis was that study skills obtained under the individualized condition would transfer to situations in which the student was not under such a method. Three designs were used to test this hypothesis.

First, the two variables in the 2 x 2 design of Figure 2 were predicted to interact in such a way that the difference between IC and

CC groups would be greater than the difference between II and CI groups on the final exam. The reasoning behind this prediction is as follows: the II and CI groups differ only in that the CI group had half as much individualized instruction as did the II group. There could be no effect of positive transfer here, since the CI group did not have individualized instruction until the second half of the semester. The difference between the IC and CC groups is also one of half-a-semester's worth of individualized instruction. However, the possibility of positive transfer occurs for the IC group, and would augment the difference between it and the CC group. Thus, if there is any transfer, IC - CC should be greater than II - CI.

The second prediction of transfer was based on the second exam, where IC should outperform CC. The material being tested on the second exam was covered under the conventional instruction for both groups; the only difference between the groups is the presence of the first half of a semester of individualized instruction for IC. If individualized instruction trains study skills, positive transfer should augment the performance of IC relative to CC on the second exam.

The third prediction of transfer was made on a chapter of reading which was assigned during the independent reading period at the university. During this time, no classes met, and students were assigned chapter 15 of the text. Final exam items which dealt with material from chapter 15 were used to analyze performance, and the prediction was that II > IC or CI > CC.

Hypothesis$_4$ (Evaluation of Mastery)

Evaluation of mastery was measured on both hourly examinations by having students place a check mark beside those multiple-choice items they were certain they answered correctly. Half an hour after the exam subjects were given photostated copies of their exams on which to estimate their accuracy. Amount of correspondence between the checks and items actually answered correctly served as the index of ability. The prediction was that the correspondence should be higher for the students from individualized instruction than for students from the conventional condition.

Other Measures

Other measures were taken to test the possibility of alternate interpretations which might be drawn from the results of the study. An attitude questionnaire was completed by students in all conditions just prior to the final exam. These questionnaires were filled out anonymously, with the exception that students were asked to indicate their section number.*

*Because students were asked to mark their section number before they answered the questionnaire, there was a possibility of experimenter demands operating. However, the writer feels this is intuitively unlikely. A more detailed consideration of the role of demand characteristics and the Hawthorne effect in terms of the study in general appears in the Discussion.

A second measure involved 50 subjects from each condition who took a recall exam immediately following the first hourly exam. The purpose of administering this test was to compare performance on a different kind of test item, namely short answer fill-in-the-blank, which required recall instead of recognition, which is required in the multiple-choice items. Five student assistants served as raters in a blind condition such that they did not know to which condition each test-taker belonged.

RESULTS

H_1 (Method)

On all measures, the data support the hypothesis that performance under individualized instruction is superior to that under conventional conditions. Mean scores on the final examination were in the predicted direction, CC scoring 56.58, CI scoring 59.96, IC scoring 60.10 and II scoring 63.54 on ninety items. Factor A, the method during the first half of the semester produced a significant main effect beyond the .0005 level of significance ($F = 12.90$, df 1, 451) and Factor B, the method during the second half of the semester produced a significant main effect at the .001 level ($F = 12.01$, df 1, 451). The lack of interaction between these two variables indicates that it doesn't matter when individualized instruction is instituted: the effects are remarkably similar in each half of the semester. Inspection of the mean performances bears this notion out: each half a semester of individualized instruction appears to produce about a 3 point difference in final examination scores.

Because total scores on the final examination do not reflect which material was learned better, scores on the first and second hourlies serve as needed support of the notion that individualized instruction produces better learning. Cells A + B (Figure 1) averaged 31.55 on the first exam, while the conventional group, cells C + D scored on the average of 27.89. Using an unweighted means analysis of variances, this difference was significant beyond the .0005 level ($F = 20.60$, df 1, 474).

Similarly, performance on the second examination indicated the same relationship. Students in the individualized condition (cell E, Figure 1) outperformed those in the conventional condition (cell H, Figure 1) with mean performance of 32.69 as opposed to 29.25. This difference is significant beyond the .0005 level ($F = 12.70$, df 1, 348).

While the significant levels are impressive on all three examinations, it should be pointed out that in each case, error was quite large. Thus the method factor accounted for only 8% of the variance in the first hourly exam scores, 6.8% in the second hourly scores, and only 5.23% (both A and B factors combined) on the final exam. The summary tables for each analysis appear in Table 1.* The Newman-Keuls tests on differences between means on the final appear in Table 3.

*While the analyses were violated with respect to the assumption of homogeneity of variance, (see Table 2) the variances observed are somewhat understandable in terms of the manipulation of the independent variable. This problem is dealt with in the Discussion section.

Table 1

Summary Tables for Analysis of Variance

Unweighted Means Solutions

1. Final Exam Scores Source of variation	SS	df	ms	F	p<
Factor A	1439.65	1	1439.65	12.90	.0005
Factor B	1332.08	1	1332.08	12.01	.001
AB Interaction	.01	1	.01	.00	
Error	50241.67	451	111.40		
2. First Hourly Exam **Source of variation**	**SS**	**df**	**ms**	**F**	**p<**
Treatment	1591.55	2	795.78	20.06	.0005
Error	18307.4	474	38.62		
3. Second Hourly Exam **Source of variation**	**SS**	**df**	**ms**	**F**	**p<**
Treatment	886.94	2	443.47	12.70	.0005
Error	1215.85	348	34.92		

Table 2

Variances

Test 1

Individualized = 34.92

Conventional = 42.92

F = 1.21 df 236,240 p <.05

Test 2

		II	CI	IC	CC
		29.75	63.39	27.58	44.58
II	29.75	--	p <.01		p <.01
CI	63.39		--	p <.01	p <.05
IC	27.58			--	p <.01
CC	44.58				--

Final Exam

		II	CI	IC	CC
		92.54	115.27	142.02	94.63
II	92.54	--	p approaches .10	p <.01	
CI	115.27		--	p approaches .10	
IC	142.02			--	p <.05
CC	94.63				--

Table 3

Newman-Keuls Tests on Differences between

Pairs of Means on the Final Examination

Adjusted for Unweighted Means Analysis

			1	2	3	4
Treatments			CC	CE	EC	EE
		Means	56.58	59.96	60.10	63.54
1	CC	56.58	---	3.38*	3.52*	6.96*
2	CE	59.96		--	.14	3.52*
3	EC	60.10			--	3.44*
4	EE	63.54				--

	r = 2	r = 3	r = 4
q (.99) (r, ∞)	3.64	4.12	4.40
$\sqrt{\text{ms error}/\tilde{n}}$ (9.99)	3.58	4.05	4.32
q (.95) (r, ∞)	2.77	3.31	3.63
$\sqrt{\text{ms error}/\tilde{n}}$ (9.45)	2.72	3.25	3.57

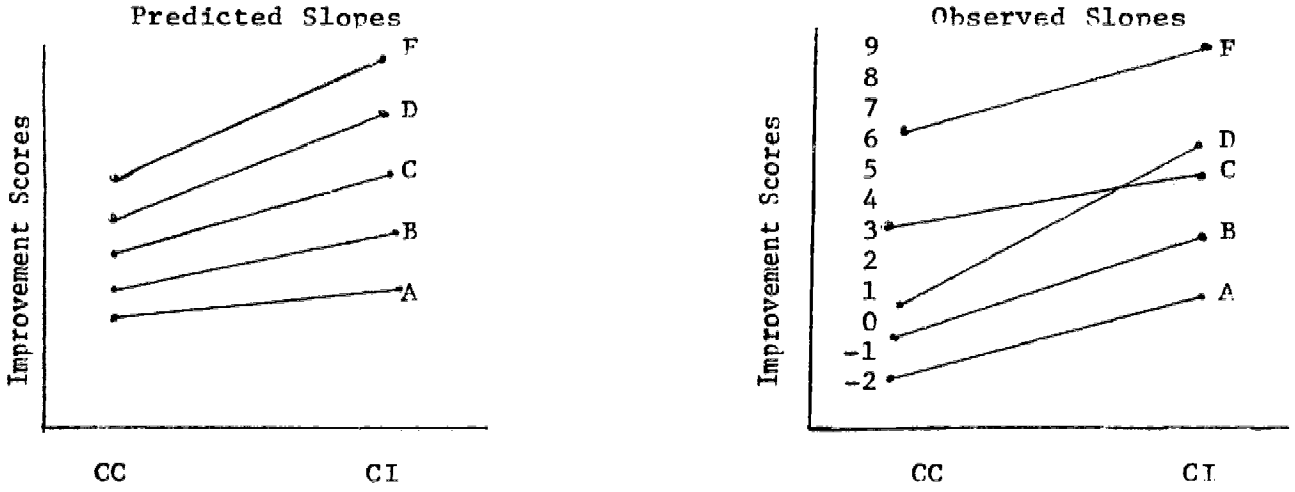
**p <.01

*p <.05

H_2 (Differential Help)

The data do not support the hypothesis that individualized instruction is more beneficial to the weaker student than it is to the stronger. While difference scores between the first and the second exam did vary inversely with letter grade in the CI group this relationship was also found in the CC group. Thus, there was a main effect for letter grade, such that the lower the letter grade received on the first exam (for both CC and CI groups) the greater the improvement on the second exam. This effect was significant beyond the .005 level ($F = 19.15$, df 1, 205). There was also a main effect for the method factor, such that improvement was always greater in the CI group than it was for the CC group ($F = 20.80$, df 1, 205 $p < .0005$). The lack of a significant interaction between the two factors, however, indicates that individualized instruction was not differentially beneficial, but rather helped students in all letter grade brackets in roughly the same manner.* The slopes of the improvement lines which were predicted, along with the observed improvement lines are presented in Table 4. The 2 x 5 analysis of variance appears in Table 6, and the post mortem Newman-Keuls analysis of difference between means in Table 7.

Table 4



*While the analysis was violated with respect to the assumption of homogeneity of variance, the differences in variance are understandable in terms of motivational differences of students who received high grades as opposed to students who received lower grades. It is thus not surprising to find that students with D's and F's showed greater variability in difference scores than did students with A's and B's (see Table 5). With the exception of slight reversals in direction of C and B students in the CI group and of B and A in the CC group, the variances uniformly decrease as a function of letter grades on the first exam.

Table 5

Variances of Difference Scores

	CC	CI
A	19.23	6.69
B	17.59	14.11
C	28.16	11.27
D	35.41	29.27
F	60.36	21.33

Table 6

2 x 5 Analysis of Variance Unweighted Means Analysis

Relative Changes in Performance CC vs. CI

Difference Scores
1st exam 2nd exam

	CC	CI
A	-1.67	.30
B	-.57	2.92
C	3.39	5.14
D	.92	5.78
F	6.35	9.5

$\tilde{n} = 21.55$

N's

	CC	CI
A	21	20
B	28	27
C	23	22
D	25	23
F	17	16

Summary Table

Source of Variation	SS	df	MS	F
Factor A	1839.51	4	459.88	19.15
Factor B	499.31	1	499.31	20.80
AB interaction	68.10	4	17.02	.71
Error	4934.28	205.5	24.01	

Table 7

Newman-Keuls Tests on Mean Difference Scores

as a Function of Grade Received on First Examination

Adjusted for Unweighted Means

Analysis

			1	2	3	4	5
Treatments			A	B	C	D	F
		Means	-.68	1.18	4.27	3.35	7.93
1	A	-.68	--	1.86	**4.95	4.03	**8.61
2	B	1.18		--	3.09	2.17	**6.75
3	C	4.27			--	.92	2.48
4	D	3.35				--	*4.58
5	F	7.93					--

	r = 2	r = 3	r = 4	r = 5
q (.99) (r, ∞)	4.02	4.64	5.02	5.29
$\sqrt{\text{ms error}/\tilde{n}}$ (8.99)	4.26	4.92	5.32	5.61
q (.95) (r, ∞)	2.92	3.58	3.96	4.23
$\sqrt{\text{ms error}/\tilde{n}}$ (9.95)	3.10	3.79	4.20	4.48

**C vs A $p<.01$
**F vs A $p<.01$
**F vs B $p<.01$
*F vs D $p<.05$

H_3 (Transfer)

The hypothesis that individualized instruction produces study skills which help the student when he is faced with typical lecture format was not borne out by the data. First of all, on the final exam, the difference between the mean performance of the CC and IC groups (3.52) was not greater than the difference in mean performance between the CI and II groups (3.58). Nor were these differences significantly different from each other.

This lack of support for the transfer hypothesis is consistent with findings on the second exam, where transfer was again not demonstrated. Recall that the prediction for transfer here was that IC would outperform CC because of IC's first half a semester of individualized instruction. However, observed means on the second exam were 29.29 for the IC group and 29.16 for the CC group. Obviously, this difference was not statistically significant.

Finally, items on the final exam which were based on the independent reading assignment also failed to show a transfer effect of individualized instruction. Table 8 indicates that the observed mean scores were not in the predicted direction, nor were they statistically different from each other.

Table 8

Means on Items for Chapter 15

from the final exam

Group	Mean Score
CC	9.64
CI	10.95
IC	9.61
II	9.91

Thus all three measures of transfer effect fail to demonstrate any improved performance due to the acquisition of study skills from individualized instruction.

H_4 (Evaluation of Mastery)

The data did support the hypothesis that students learn to evaluate their mastery while under individualized instruction. On the first exam 37 students from the individualized condition were accurate in predicting correct items 85.1% of the time, while 38 students from the conventional condition were accurate 79.5% of the time. This difference was significant beyond the .0005 level ($t = 6.37$, df 73). On the second exam 52 individualized students predicted with 87% accuracy as compared with 50

conventionalized students who predicted with 82%. This difference was also highly significant ($t = 6.98$, df 100, $p < .0001$).

Other Measures

In order to test the possibility that superior performance of students under individualized instruction was due to their becoming good multiple-choice test-takers, as opposed to their learning the subject matter better, a 45 item recall exam was administered to 50 Ss from each condition immediately following the first exam. The raters (five student assistants) grades the papers blind, and the reliability of their scoring was high, inter-judge correlations ranging from .84-.95. This agreement was achieved by deciding beforehand the range of acceptable answers.

The mean performance of students from the individualized group was 30.06, which was significantly higher than that of students from the conventional group, where the mean was 25.71 ($t = 2.47$, df 97, $p < .02$). Table 9 indicates the summary data for this recall examination.

Table 9

Summary Data for t-test on

Recall Examination-Adjusted for

Unweighted Means Analysis

Group	N	Mean	SD	t	df	p<
I	47	30.06	9.01	2.47	97	.02
C	52	25.71	9.53			

Results from the attitude questionnaire are dealt with in the discussion section as implications from it are treated.

CONCLUSIONS AND RECOMMENDATIONS

(Discussion)

The hypothesis that material would be better mastered under the individualized condition than under the conventional condition was supported by the present study. What explanations exist for such an effect?

The Effect as a Design Artifact

Although the data appear to support the view that individualized instruction is superior, it is necessary to examine the possibilities that the effects are due to artifacts of the experimental design. Several considerations are relevant.

First, there is the possibility that the superior performance of students receiving individualized instruction reflects their improved

skills in taking multiple choice tests rather than their greater mastery of the subject matter. Previous research cited above does not support this possibility, as it has been repeatedly demonstrated that superior performance generalizes across different kinds of tests items (Sheppard & MacDermot, 1970; Alba & Pennypacker, 1971; Born et al., 1971). Evidence from the recall exam conducted in the presented study provides additional support for the conclusion that individualized instruction produces superior learning of course content, as opposed to taking multiple choice tests, since students from the individualized condition also outperformed their control counterparts on the recall exam.

A second possibility is that individualized instruction produced superior results simply because the instructor was poor, and students in the individualized group had less exposure to him. The attitude questionnaire does not support this explanation, however, as the majority of students in all conditions report the instructor's classroom performance as "excellent" or "good" as opposed to "fair" or "poor". Similarly, lectures were rated as "very well organized" or "well organized" more frequently than "poorly organized" or "very poorly organized" by all groups, with no significant differences among them. Similar reactions occurred to a question concerning the interest value of the instructor's lectures. A majority of all students rated the lectures "very interesting" or "interesting" rather than "boring" or "very boring" with no significant differences occurring between the various groups (see appendix, questions 1, 2, and 5).

One might argue that the effect was due to the dropping of the weakest students from the experimental condition. Such a result was noted in one study (Born et al., 1971) although another study reported that grade point averages of students who dropped the individualized course were comparable to those of students who dropped the conventional course (Sheppard & MacDermot, 1970). Since over 90% of the students enrolled in the courses studied in the present research were freshman, it was not possible to check their grade point averages before they were enrolled in the course. However, approximately the same numbers of students dropped each condition, as Table 10 indicates.

Table 10

N at Beginning and End of Semester by Groups

	II	IC	CI	CC
Start	140	140	140	140
Finish	118	111	113	116

Students dropped the individualized condition somewhat earlier than did those who dropped the conventional condition, but this was expected as students in the individualized group received weekly information beginning with the very first week regarding the probability of their passing the course. Students in the conventional condition waited to drop until after the first examination, which was their first opportunit to evaluate their progress in the course.

Although arrangements were made with the registrar to employ the Doomsday Contingency wherever needed, in actuality no one was forced to leave the course. A "fudge factor" was employed, allowing students to pass with 7 out of 10 on the fourth quiz, if they had attended all the sessions, and allowing one student to get by with 6 out of 10 because of a special excuse. The effect of this leniency was probably to decrease the size of the difference between individualized and conventional groups performances, since the weakest students were allowed to remain in the course and thus pull the overall performance of the individualized instruction group down. However, since the researcher was interested in testing the effectiveness of the method, not the rate of attrition it produced, this was a necessary procedure.

Finally, how much faith can be placed in the effect as it was measured in this study? While the differences in performance are not overwhelming, they are consistent. Examination of mean scores indicates that on each of the exams, there was approximately a three item increase for each half a semester of individualized instruction. This remarkable consistency indicates that the critical factor is how much individualized instruction a student receives, rather than when he receives it.

Could the results be due to chance fluctuations in measurement as a result of unreliable examinations? Split-half correlation measures on each of the exams indicates that they are comparable in reliability to those of area examinations on the Graduate Record Examinations. Of the 19 achievement tests of the G.R.E. studied, 4 yielded coefficients of .80 or better, and the remaining 15 yielded coefficients of .90 or better. These reliability coefficients were described as satisfactory (Seashore, as reported in Buros, 1959). In the present study, the first exam yielded a coefficient of .78 and the final exam a coefficient of .87, yielding standard errors of measurements of 3.47 and 4.56 respectively (data were not available for the second hourly exam). Thus, it appears that the reliability of the exams used in this study are typical of achievements tests, and also appear to be satisfactory.

The Hawthorne Effect

Perhaps the results of this study can be explained in terms of the "Hawthorne effect", named after the now classic studies at the Hawthorne plant of the General Electric Company (Roethlisberger & Dickinson, 1940). The investigators of this study had difficulty drawing conclusions from their research on the effects of illumination, rest-periods, wage incentives, and other variables on work output. It was not clear whether the employees increased performance was due to these variable which the experimenters intentionally manipulated, or to the unintentional effects such as competition induced between the experimental and control groups or the desire to perform well under various experimental conditions. Thus the Hawthorne effect has been defined as an increase in performance occurring from special attention received while participating in an experiment, and has become an important variable to control for in research with human subjects.

However, it has been suggested that the Hawthorne effect is probably not due to a single variable, and deserves attention in its own

right (Sommer, 1968). The problem as it relates to the present research becomes one of definition. If one means by the Hawthorne effect, the fact that performance increases because more attention is given, then the phenomenon is itself part of the experimental variable, for individualized instruction is seen as arranging the large lecture section so that students can receive more individualized attention. However, if one defines the Hawthorne effect as increased performance due to attention resulting from being subjects in an experiment, then it becomes an extraneous variable and would be a concern for which an experimenter should control. Several observations are relevant.

In an effort to equate the Hawthorne effect for all sections, the instructor announced at the beginning of the semester that all sections were part of an experimental design, and told each group that it was the experimental group. However, some students identified themselves as belonging to "the experimental condition" or the "control group" and these descriptions were usually made by students in the individualized and conventional conditions, respectively. The fact that the experimental manipulations were obvious is not surprising in that two very different teaching methods were delivered to students who studied topics concerning experimental design and control concerns during the second week of the semester. However, was this knowledge enough to account for the increased performance under individualized instruction?

Research on the Hawthorne effect in the classroom would indicate not (Johnston & Foley, 1969). In this study, three groups of students were given different instructions regarding the same method of instruction. Specifically, groups of students were told that they were either a) filling some time (time filler), b) trying a teaching method of undetermined value (experiment) or c) using a teaching method of demonstrated excellence (placebo). Scores on a multiple choice quiz administered after the treatments indicated that the placebo group performed significantly better than the experiment and time-filler groups, but performance differences between the latter two were negligible. Thus, if the Hawthorne effect was operating, one would expect the experimental group to be different than the time-filler group. The individualized instruction group in the present study is analogous to the experimental group in the study just described, as the students in it were told that they were being instructed under a method of undetermined value.

More direct evidence from the present research also does not support an explanation in terms of the Hawthorne effect when it refers to experimenter demands. As the IC group moved from the individualized instruction method to the conventional method, its performance decreased. Yet analogous changes in the Hawthorne studies were followed by increases in performance. It is interesting to note here that the attitude questionnaire administered just before the final exam indicates that the IC group had significantly more feelings of being used as a "guinea pig" than any of the other groups (see appendix, question 32). Thus, while the IC group received the same amount of attention and experimental manipulation as the CI group, those students taken off of what they felt was the more desirable condition reported more dissatisfaction. Such observations would seem to be more a function of the particular manipulation than

manipulation per se.*

Theoretical Rationals

It appears that the effects observed in this study are not artifacts of the experimental design and that individualized instruction is superior to conventional large class procedures. Why is this so?

The most frequently cited explanation of the efficacy of individualized instruction has been in terms of an operant conditioning model (Keller, 1968; Ferster, 1968; Malott & Svinicki, 1968; Domjan & DuNann, 1969; Johnston & Pennypacker, 1971). Partiality to this kind of explanation is not surprising in view of the fact that the method grew out of formulations by educators working within the operant conditioning camp. According to this view, students in large classes are expected to perform on infrequent hourly exams covering large amounts of material, although they have not been required to emit consistent study behavior beforehand. Individualized instruction is seen as the rearrangement of the classroom situation such that reinforcement scheduling is more appropriate for producing consistent study behavior. It is assumed that students enjoy receiving good grades, and because the quizzes cover small units of material and the student has several opportunities to pass them, it is a relatively easy task to get a good grade on a quiz. The frequency with which the quizzes are scheduled allows the student to shape his behavior before the hourly exams, as the desired responses are clearly specified and the students responses are evaluated relative to the criterion. In addition to the frequent reinforcement in the form of good grades, individualized instruction also allows the delivery of an additional kind of reinforcement, that of social reinforcement, through the use of student assistants. Students are given the opportunity to verbalize ideas and questions and engage in discussion of the material, while a fellow human being actively responds.

This use of human interaction also suggests another explanation for the superiority of individualized instruction: That it is effective because it rectifies some of the maladies of the depersonalized large lecture classroom. In other words, the method works because it alleviates the aversive effects of large sections, which Jensen (1968) has lucidly described as:

> . ...reluctance to attend class, "sour" facial expressions, apathy and studied indifference while in class, eagerness to leave class and joy if a class is cancelled--these typical students behaviors all suggest that for the majority of students the large lecture class is an aversive event, i.e. an unpleasantness to be

*While the differences in attitude and performance of the CI versus the IC groups could be attributed to different demandscharacteristics, it seems difficult to separate such effects from the independent variable itself. Thus, it is felt that the likelihood is small that experimenter demands produces an effect which was independent of those "demand characteristics" produced by individualized instruction itself.

avoided, quickly terminated, or grudgingly endured rather than a rewarding experience to be sought, prolonged, and savored (pp.1-2).

Thus, the explanation offered is that individualized instruction is effective because it reduces the aversive effects of large lecture sections by reducing depersonalization, independent of whether it produces better learning through more consistent study behavior.

Another explanation is one which involves the facilitory effects of information feedback. The regularly scheduled delivery of information regarding the performance of each student as he progresses through the course allows him to evaluate the efficacy of his study behavior. This is an analogous explanation to the shaping discussed above within the operant model, but with an emphasis on the cognition, or understanding of the student as he processes information relative to his goals.

Consistent with the above explanations, but somewhat different from them is another in terms of the alleviation of excessive anxiety, and the detrimental effects of too much anxiety on test performance. The inadequate reinforcement scheduling provided by courses including only mid-term and final exams permits students the possibility of falling into the common, but dangerous, loaf-and-cram syndrome, which in turn results in excessive anxiety that may interfere with rather than facilitate study behavior (McKeachie, 1968; Jensen, 1969). It has been well established that anxiety facilitates task performance when the task is simple and hinder performance when the task is complex (Taylor, 1957; Farber & Spence, 1953) or that it facilitates when the response called for is dominant in the hierarchy of responses, but hinders when the response called for is subdominant (Palermo, 1957; Ruebush, 1963). Thus while anxiety may facilitate a good student's performance (since his response is dominant and the task of answering questions presumably simple) it is reasonable to expect that a large number of students who cram before exams are hindered by anxiety since they are faced with a complex task of trying to emit responses which are sub-dominant.

This reasoning suggests that the variability of scores for students in the conventional condition should be greater than that for the individualized condition. While this effect was clearly observed on the first examination, the relationship is not a clean one on the other measures (see Table 2). While the CC group was significantly greater than II on the second exam, the CI and IC groups' variability is difficult to explain. CI was significantly greater than CC and IC was similar to II. On the final exam, the relationship is even less clear, where II and CC groups were lower than CI, which was lower than IC.

While the explanations in terms of an operant conditioning model, more personalization, information feedback, and lack of interfering anxiety all seem intuitively likely, still another explanation is that individualized instruction provides a policed situation that induces study through fear of failing and/or having a student assistant breathing down the student's neck. However, questionnaire data strongly suggests that the course was anything but a purely aversive experience. Students

in both the individualized and conventional conditions preferred the individualized condition, and students in the individualized condition wished that other courses were run with the individualized method. Furthermore, they liked the course better than did students in the conventional condition, found it less difficult, and more interesting. They also perceived the grading as more lenient, in spite of the fact that performance on quizzes did not count towards grades (see appendix, questions 31, 44, 25, 22, 23, and 20).

These responses are surprisingly positive, especially in light of the fact that credit was not given for quiz performance, as it was in many other studies, and that the original instructions to the students were delivered in a somewhat negative fashion, with undue emphasis placed on the "Doomsday Contingency." However, these attitudinal measures do agree with those taken by previous investigators who report positive attitudes of students experiencing the individualized course (Domjam & DuNann, 1969; Johnston & Pennypacker, 1971; Sheppard & MacDermot, 1970; McMichael & Corey, 1969; Born et al., 1971). While it is understandable that students feared failing quizzes and still liked the course, the aversiveness of the large classroom appears to be diminished in the individualized condition, and hence the method appears valuable in terms of affective as well as cognitive goals.

In summary, there are several explanations of the effects of individualized instruction. The research was not designed to test one explanation over another. However, the attitudinal measure suggests that an explanation in terms of avoidance conditioning is least satisfactory.

Differential Help

The hypothesis suggested by Born et al. (1971) that individualized instruction is more beneficial to weaker students than it is to stronger students was not supported by the data. This finding stands in sharp contrast to that of Born et al. and is interpretable in terms of the different designs used to test the hypothesis. Born et al. noted better improvement for weaker students as opposed to stronger students as they changed from a conventional to an individualized condition. This finding is explainable in terms of differential motivation for high-scoring and low-scoring students, as well as a statistical effect of regression towards the mean. Since this same relationship of grades and improvement scores was also noted in the CC group in the present study, the explanation by Born et al. of it being due to individualized instruction appears unwarranted. The lack of interaction of the method factor and the ability level factor in the present study indicates that individualized instruction did not differentially help weaker versus stronger students. Thus, the conclusion drawn by Born et al. appears to be due to an inadequate design which did not allow comparison to a necessary baseline.

Transfer of Training

Data relevant to the transfer hypothesis indicate that students do not profit from their experience under individualized instruction when

later faced with the lecture situation, or when pursuing an independent study experience. This finding contradicts the findings of a previous study on transfer cited above (Sassaranth & Garverick, 1965). One obvious explanation for the discrepancy is that the measure of transfer in the earlier study consisted of different items on the same material, while the present study used different items on different material.

Another reason for the failure of transfer to be demonstrated could be that the attitudes of the IC group, which were more negative regarding their being used as "guinea pigs", interfered with their performance, which might have been superior to the CC group on the second exam had their attitudes been similar. Thus, it is possible that the negative effects of changing from a more desirable condition interfered with academic performance and negated any effects of transfer.

It appears then, that the method has a short term effect which is of immediate use in mastering textbook material, rather than in training study skills for long-term use. The literature on college study courses would not contradict this conclusion, since in these cases, a much more intensive and direct attempt is made to change study behavior. Whether or not such a goal can be met within the limited confines of a single semester devoted to content material remains to be answered, but does not appear likely.

Evaluation of Mastery

The hypothesis that students would learn to evaluate their mastery of material more adequately when under individualized instruction was supported by data from both hourly exams. This highly important finding provides additional support for the superiority of individualized instruction and helps explain why students experiencing individualized instruction perceived grading as more lenient, despite the fact that performance on quizzes did not count towards grades. Thus it seems reasonable that these students saw grading as more lenient because they were acquiring a greater ability to evaluate their performance, and because of this grading was seen as fairer.

Pragmatic Concerns

Although this study indicates that the individualized method is of more immediate value in increasing the amount of learning of textbook material, it is still meaningful to ask whether the size of the effect is pragmatically significant. Is a 6 point increase on a 90 item test worth the time, energy, and money needed to set up such a course? There is not clear-cut answer to this question, but some comments are relevant.

First, once the course is set up, the instructor spends less time in class than he does with the conventional lecture procedure. Obviously, the real expenses come before the course begins, in selecting and training student assistants, generating quizzes, and the working out of administrative and other details. The course described in this study was set up in approximately 100 hours, 50 spent in selecting student assistants, 20 in generating quizzes from an already collected item pool, and 30 in working out administrative details. Each week a department

secretary typed and printed four ten-item quizzes, and a graduate student assistant helped in the administrative tasks of coordinating record-keeping done by the student assistants and handling other problems. With the moderate additions of some hours spent before the course, plus some secretarial and graduate assistant service, the course involved less class time on the part of the instructor during the semester than did the conventional lecture method. The instructor's shortened classroom time was made up for by the fact that he spent a few hours each week working with individual student assistants. Therefore, a total time devoted to the method by the instructor would roughly equal that spent with the conventional lecture method.

Another practical concern pertains to the question of whether the individualized method is effective only because it forces students to spend so much time studying for quizzes that their performance in other courses suffered. Questionnaire data suggests that this was not the case. Students taught under individualized instruction reported that the course was less difficult than other courses taken at the university and that work habits in other courses did not change as a result of taking the course (see appendix, questions 22 and 37).

In summary, conclusions drawn from this study are that the effects of individualized instruction are immediate, reliable, and not especially large. Students prefer the individualized condition over the conventional one. However, individualized instruction does not produce the overwhelming effects often described by eager enthusiasts, nor is it clear that the effects it does produce are justifiable in terms of the added expenses of secretarial and graduate assistance listed above.

Recommendations

Given that individualized instruction is moderately successful in improving students' understanding of subject matter as well as producing positive attitudes about the course, future research on the method appears warranted. What was studied in this project is a global factor consisting of many variables which probably differ in importance. One of the most obvious questions which needs to be answered concerns the relative effects of social interaction, feedback, structure, positive reinforcement, and the Doomsday Contingency in producing improved performance. Future research might begin with testing out some of these factors through the use of more complex multivariate designs. The study of the relative effectiveness of the different variables might have important pragmatic value for future instruction. For instance, why train student assistants if the effect is due primarily to the use of weekly quizzes which could be handled in larger groups, and possible through the use of a computer? Or, similarily, why use quizzes at all if the more important gains are made through the use of social interaction?

However, it is important to note that before separation of factors is attempted, a more fruitful next step might be to try to increase the size of the effect, since the separating of variables on the delicate effect observed in this research might eliminate it. In the present research, significance undoubtedly rested in part upon the use of a large N; future research which uses smaller samples might not yield

significance at all.

The study of individual differences would be valuable in this respect. The finding that the method had its greatest effect on students with lower grades is a case in point. Perhaps other important student characteristics could be identified. It would seem that a study habits measure might be a likely candidate for such a purpose. Additionally, personality variables such as needs for achievement and for affiliation, internal-external control, or others, might also yield valuable data, not only in terms of maximizing the effect, but also in shedding some light on theoretical concerns. While it is unlikely that university registration processes would be able to use such information in the near future, study of these factors would give a much better understanding of what kind of phenomenon is actually occurring.

REFERENCES

Alba and Pennypacker, H. S. (title unavailable). Unpublished manuscript, Univ. of Florida, 1971.

Born, D. G., Gledhill, S. M., and Davis, M. L. Examination performance and number of student withdrawals in lecture-discussion and personalized instruction courses. Unpublished manuscript, University of Utah, 1971.

Day, H. G. Report of the faculty committee on facilities for large group instruction. Indiana University, 1966. Cited by D. D. Jensen. An effective, and humane system for large class instruction. Paper presented at the meeting of the Midwest Psychological Association, Chicago, May, 1969.

Di Lorenzo, L. T. Discriminating effects of a college study course, Journal of Educational Research, 1964, 57, 472-475.

Domjam, M. P. and Du Nann, D. H. Contingency management in an introductory laboratory course. Midwest Education Review, 1969, 1, 35-46.

Entwistle, D. Evaluation of study skills courses: a review. Journal of Educational Research, 1960, 243-251.

Farber, I. E., and Spence, K. W. Complex learning and conditioning as a function of anxiety. Journal of Experimental Psychology, 1953, 45, 120-125.

Ferster, C. B. Individualized instruction in a large introductory psychology course. Psychological Record, 1968, 18, 521-532.

Jensen, D. D. An efficient, effective and humane system of large class instruction. Paper presented at the meeting of the Midwest Psychological Association, Chicago, May, 1969.

Johnston, H. H. and Foley, J. M. Some effects of placebo and experiment conditions in research on methods of teaching. Journal of Educational Psychology, 1969, 60, 6-10.

Johnston, J. M. and Pennypacker, H. S. A behavioral approach to college teaching. American Psychologist, 1971, 26, 219-244.

Keller, F. S. Good-bye teacher. Journal of Applied Behavior Analysis, 1968, 1, 78-89.

Malott, R. W. and Svinicki, J. G. Contingency management in an introductory psychology course for 1000 students. Mimeographed manuscript, Western Michigan University, 1968.

McKeachie, W. J. Psychology at age 75--the psychology teacher comes into his own. American Psychologist, 1968, 23, 551-557.

McMichael, J. S. and Corey, J. R. Contingency Management in an introductory psychology course produces better learning. Journal of Applied Behavior Analysis, 1969, 2, 79-83.

Melarango, R. J. Two methods for adapting self-instructional materials to individualized differences. Journal of Educational Psychology, 1967, 58, 327-333.

Palermo, D. S. Proactive interference and facilitation as a function of amount of training and stress. Journal of Experimental Psychology, 1957, 53, 293-296.

Rentel, M. Improving study skills, Journal of Reading, 1966, 9, 343-348.

Ripple, R. E., Millman, J., and Slock, M. D. Learner characteristics and instructional mode: a search for disordinal interactions. Journal of Educational Psychology, 1969, 60, 113-120.

Roderick, M. and Anderson, R. C. Programmed introduction to psychology versus text. Journal of Educational Psychology, 1968, 59, 381-387.

Roethlisberger, F. J. and Dickinson, W. J. Management and the Worker, 1940, Cambridge, Mass.: Harvard University Press.

Ruebush, B. K. Anxiety. In H. W. Stevenson (Ed.), Child Psychology 62 Yearbook, National Social Studies Education, Part I, 1963, Chicago: University of Chicago.

Sassaranth, J. M. and Carverick, C. M. Effects of differential feedback from examinations on retention and transfer. Journal of Educational Psychology, 1965, 56, 259-263.

Sheppard, W. C. and MacDermot, H. G. Design and evaluation of a programmed course in introductory psychology. Journal of Applied Behavior Analysis, 1970, 3, 5-11.

Sommer, R. Hawthorne dogma. Psychological Bulletin, 1968, 70, 592-595.

Stone, G. R. The tra..ng function of examinations: retest performance as a function of the amount and kind of critique information. Research Report #AFPT-RC-TN-55-8, 1955, U.S.A.F. Personnel Training Research Center, Lackland A.F.B.

Taylor, J. A. The relationship of anxiety to the conditioned eyelid response. Journal of Experimental Psychology, 1951, 41, 81-92.

Time, "Education" 1971, v. 97, #24 June 14, 1971, p. 52.

Van Dam, R. H. Programmed lectures as a means of increasing student achievement and involvement in large lecture classes. Dissertation Abstracts, 1968, 29 (2-A), 522-523.

Volin, B. A new sociology of learning. Journal of Emotional Education, 1967, 7, 28-32.

APPENDIX

Questionnaire data and

corresponding chi square analyses when appropriate

#1. Question: Lectures by the instructor were

a. percent responding by groups

	CC	CI	IC	II
very well organized	8	7	4	4
well organized	44	47	39	41
O.K.	43	40	48	44
poorly organized	3	5	8	8
very poorly organized	1	0	2	3

#2. Question: Lectures by the instructor were

a. percent responding by groups

	CC	CI	IC	II
very interesting	7	7	4	5
interesting	50	48	47	45
so-so	34	35	29	37
boring	5	8	16	9
very boring	3	2	5	3

#5. Question: Overall I rate the instructor's classroom performance as

a. responses: Percent responding by groups

	CC	CI	IC	II
excellent	15	13	8	13
good	46	44	42	43
satisfactory	26	32	28	30
fair	8	9	17	8
poor	6	3	5	6
Total	100	100	100	100

#20. Question: grading was

a. percentage responding by groups:

	CC	CI	IC	II
too lenient	1	0	0	2
lenient	2	5	12	7
fair	68	72	67	69
strict	22	17	15	16
too strict	7	5	5	5

b. X^2--Frequency of responses by groups

expected/observed

	CC	II
lenient	6.53 / 3	5.47 / 9
strict	30.47 / 34	25.53 / 22

$X^2 = 5.09$

df = 1

$p < .05$

#22. Question: Relative to other courses I've taken at this university, this course is

a. responses: percentage responding by groups

		CC	CI	IC	II
Cells Combined For Analysis	much more difficult	17	7	10	9
	more difficult	33	42	29	35
	of about equal difficulty	41	38	41	37
Cells Combined For Analysis	less difficult	9	12	12	15
	much less difficult	0	1	2	2

B. X^2 analysis: frequency of responses

expected/observed

	CC	II
more difficult	56.89 / 51	53.10 / 59
less difficult	18.10 / 24	16.40 / 11

$X^2 = 5.25$

df = 1

$p < .05$

#23. Question: relative to other courses I've taken at this university, this course is

a. percentage responding by groups

	CC	CI	IC	II
much more interesting	8	10	15	8
more interesting	36	44	38	47
of about equal interest	35	31	29	33
less interesting	14	13	16	8
much less interesting	7	3	3	3

b. X^2--frequency of response by groups

expected/observed

	CC	II
more interesting	56.89 / 51	53.10 / 59
less interesting	18.10 / 24	16.90 / 11

$x^2 = 5.25$

df = 1

$p < .05$

#25. Question: relative to the other co..ses I've taken (or am taking) at this university this course was

a. percentage responding by groups:

	CC	CI	IC	II
much better	5	7	12	12
better	28	36	32	35
about the same	40	38	37	42
worse	22	17	16	7
much worse	4	3	4	3

b. X^2--frequency of response

expected/observed

	CC	II
better	47.92 / 39	41.07 / 50
worse	22.08 / 31	18.92 / 10

$X^2 = 9.93$

df = 1

$p < .01$

#31. Question: I would have preferred to have been in the

[lecture / quiz feedback] rather than the

[quiz feedback / lecture] condition

a. Preference for opposite condition percent responding:

	II	CC
Yes	3	48
No	81	20
Don't care	5	13
Don't know	8	17

b. X^2--frequency of report

expected/observed

	II	CC
yes	31.26 / 3	27.74 / 56
no	57.74 / 86	51.36 / 23

$X^2 = 83.74$

df = 1

$p < .001$

#32. Question: As a student in psychology 401 you were also a subject in an experiment in teaching. Did this make you feel used? That is, did you feel like a "guinea pig"?

a. responses: Percent Responding by groups

	CC	CI	IC	II
yes	16	19	32	16
somewhat	33	32	31	32
no	38	47	37	46
omit	3	2	0	6
Total	100	100	100	100

b. X^2 analysis--Number of Responses by Groups

expected/observed

	CC	CI	IC	II
Yes	23.79 / 19	23.48 / 21	23.79 / 35	20.94 / 17
No	51.21 / 56	50.52 / 53	51.21 / 40	45.06 / 49

$X^2 = 9.2$

df = 3

$p < .01$

	CC	CI	CC
Yes	19.88 / 19	19.62 / 21	19.88 / 19
No	55.12 / 56	54.38 / 53	55.12 / 56

$X^2 = .2$

df = 2

N.S.

#37. Question: As a result of my experience in Psychology 401 my study skills and work habits in other courses

a. responses: percentage responding in each group

	CC	CI	IC	II
Improved	10	9	13	13
got worse	9	6	9	10
were unchanged	78	83	78	74

#44. Question: I wish the other courses I am taking this semester had used the quiz-feedback method

a. percentage responding by groups:

	CI	IC	II
yes	59	48	55
no	28	36	30
doesn't matter	10	16	15